Dual character concepts in social cognition: Commitments and the normative dimension of conceptual representation*

Guillermo Del Pinal[*] and Kevin Reuter[**]

[*] *Center for General Linguistics (ZAS), Berlin*
[**] *Institute of Philosophy, University of Bern*

forthcoming in *Cognitive Science*

September 8, 2016

Abstract

The concepts expressed by social role terms such as *artist* and *scientist* are unique in that they seem to allow two independent criteria for categorization, one of which is inherently normative (Knobe et al., 2013). This paper presents and tests an account of the content and structure of the normative dimension of these 'dual character concepts'. Experiment 1 suggests that the normative dimension of a social role concept represents the *commitment* to fulfill the idealised basic function associated with the role. Background information can affect which basic function is associated with each social role. However, Experiment 2 indicates that the normative dimension always represents the relevant commitment as an end in itself. We argue that social role concepts represent the commitments to basic functions because that information is crucial to predict the future social roles and role-dependent behavior of others.

Keywords: dual character concepts; categorization; concepts; prototypes; commitments; social cognition; normative dimension of conceptual representation.
Word count: 9,341

*We would like to thank Joshua Knobe for his invaluable support for this project. We are also grateful to Luca Barlassina, James Hampton, Martin Jönsson, Lars Kuchinke, Alex Madva, Gregory Murphy, Eleonore Neufeld, Sandeep Prasada, Shannon Spaulding and three anonymous referees for comments on this paper. This work was presented at conferences in Bern, Nottingham, Osnabrück, Pasadena and Reading. We would like to thank the audiences for their feedback. The research for this paper was supported by the Alexander von Humboldt Foundation.

1 Introduction

In a recent paper, Knobe, Prasada, and Newman (2013) present a series of experiments designed to show that the concepts expressed by terms such as *artist* and *scientist* have two independent criteria for categorization, one of which is inherently normative. They call this unique class of concepts, paradigmatically expressed by social role terms, 'dual character concepts' (DCCs). The study of DCCs promises to shed light on the nature of our conceptual structures, and has important implications for topics in social cognition such as normative generics and gender biases (Leslie, 2015a,b; Reuter and Del Pinal, 2016). In this article, we present an account of the content and structure of the normative dimension of DCCs, and test it in a series of experiments.[1]

To illustrate the basic properties of DCCs consider the following scenarios. Suppose John has two biological children, whom he materially provides with all their needs, including food, good schools, and nice toys. However, John does not think doing so is his duty. He is only a good father because he thinks that will advance his career as a politician. Is John really a father? In response, consider whether you agree with (1-a) and (1-b):

(1) a. There is a sense in which John is clearly a father.
b. However, if you think about what it really means to be a father, you would have to say that John is not a true father after all.

Now, imagine that John's career foundered and that he abandoned his two kids. Fortunately, John's brother, Mark, who has no biological children, decided to care for the kids. Mark is not as materially successful as John. Sometimes he can't provide them with some of their needs. Still, Mark really loves the kids and works very hard to support them. Is Mark really a father? In response, consider whether you agree with (2-a) and (2-b):

(2) a. There is a sense in which Mark is clearly not a father.
b. However, if you think about what it really means to be a father, you would have to say that Mark is a true father after all.

The experiments presented by Knobe et al. (2013) suggest that most of us would accept (1-a)-(1-b) and (2-a)-(2-b). They also show that this basic pattern generalizes to other social role terms such as *artist*, *scientist*, and *friend*.

To fully see what is special about DCCs, it is important to note two points. First, as emphasised by Leslie (2015a), note that the acceptability pattern just illustrated cannot be explained as a case of standard raising. One might think

[1] We should clarify at the outset that although—for continuity with previous literature—we will use the term 'dual character concepts' and assume that social role terms express DCCs, we do not thereby presuppose any position regarding the general conceptual structure of social role terms, esp., regarding the number of dimensions that they encode. Our only assumption is that DCCs include, among other dimensions, a normative dimension, and our task here is to investigate the content and structure of this normative dimension, as it is manifested in social role concepts. We discuss the general conceptual structure of social roles in the 'General Discussion'.

that John does not count as a *true father* simply because he fails to meet the highest standards of fatherhood. However, Mark does count as a *true father*, although he is not a normal or standard father. Since in paradigmatic cases of standard raising—e.g., in *perfect athlete*—a 'highly rated x' is still an x, the pattern captured by (2-a)-(2-b) suggests a more complex interaction between linguistic context and the conceptual structure of DCCs. Secondly, social role terms such as *welder*, *bartender* and *bus driver* obtain lower ratings in cases parallel to (1-b) and (2-b). In other words, there are differences amongst social role terms in how naturally they can be accommodated into patterns analogous to (1-a)-(1-b) and (2-a)-(2-b).

To explain these patterns, Knobe et al. argue that DCCs have a unique internal structure. As with concepts of other kinds, we can use the concrete features associated with a DCC to categorise entities and events. What is distinctive of DCCs, however, is that they also include a normative dimension that is relevant for categorization. As illustrated in the case of *father*, the normative dimension can be dissociated from other concrete features of the class and used as the sole categorization criterion. In particular, modifiers such as *true* in colloquial expressions such as *John is not a true father*, and *real* in expressions such as *John is not a real man*, seem to operate (not necessarily exclusively) on this normative dimension. What, precisely, is the content and structure of this normative dimension?

Knobe et al. (2013) make some important points that bear on this question. The first is that the normative dimension has to do with abstract values. What they mean by this is revealed by their examples: for *scientist* the abstract value is something like the quest for impartial truth, and for *artist* something like creating works of deep aesthetic value. The second is that the normative dimension of a class must be distinguished from (at least some notions of) its typical function. For example, while the normative dimension interacts with modifiers such as *true*, it does not seem to directly do so with modifiers such as *good*, in the sense of 'useful' or 'efficient'. Experiment 1 of Knobe et al. (2013) shows that most social role terms are equally and highly acceptable when modified by *good*, whereas their acceptability varies substantially when they are modified by *true* (see the Preliminary Study below).

In an interesting discussion which builds on the results of Knobe et al., Leslie (2015a) proposes a specific view of the content of the normative dimension, focusing, as we do here, on the case of social role terms. Leslie agrees that the normative dimension of DCCs does not represent the usual or even typical function of the corresponding social roles. For example, the normative dimension of *scientist* does not represent superficial functions such as gathering data, looking into microscopes, etc. What is represented is more fundamental: it is more like an idealisation of the basic function of the role. In the case of *scientist*, this would be something like building theories that are responsive to empirical evidence, among other things. On this view, we can say that a scientist that, even if in some sense skillfull, completely resists empirical demands to modify or abandon his theory is not a true scientist because he does not satisfy the idealised basic function of a scientist.

Leslie's proposal can be interpreted in at least two ways. One is to assume that the normative dimension just represents the idealised basic function. On this view, the normative dimension of *father* would directly represent something like providing for the needs of the family, and that of *scientist* would represent something like the search for empirically constrained theories. We will argue that this is not the best way to implement Leslie's proposal. It is in tension with the observation, mentioned above, according to which someone who tries hard but often fails to fulfill the basic function of a father or scientist can still count as a true father or scientist. Experiment 1 below confirms this intuition, namely, that the commitment to fulfill the relevant role is crucial to satisfy the normative dimension. In addition, both Experiments 1 and 2 show that skillfully or efficiently fulfilling the basic function is not enough to fully satisfy the normative dimension.

In our view, a better way to interpret Leslie's proposal is to hold that, to satisfy the normative dimension, what matters most is not whether someone actually fulfills the basic function to some non-trivial degree, but rather whether someone is *committed* to fulfill it. Consider again the example of a non-biological father. From this perspective, what made us agree that Mark is a *true father* despite not being a biological father and often failing to fully provide for the kids, is his sincere and constant commitment to care for them. We should distinguish, then, between the property of actually fulfilling the basic function of a social role from that of being *committed* to fulfill it. The most important property for satisfying the normative dimension is having the commitment to the basic function. Experiments 1 and 2 below test this view of the way in which commitments and basic functions are represented in the normative dimension.

To be clear, our claim is not that evidence that an entity fulfills the basic function of a DCC does not matter for satisfying its normative dimension. After all, that someone is a skillful scientist is often a reliable sign that s/he is committed to being a scientist. Our claim is that being committed to fulfill the basic function is more important to satisfy the normative dimension of DCCs than simply being able to fulfill that function. Furthermore, we do not assume that social role terms are associated with fixed basic functions across all contexts. Indeed, as shown in Experiment 2, background information can influence which basic function we associate with a social role. For example, in some contexts, we can represent artists as aiming to create deep—if disturbing—works of art, while in other contexts we can represent them as aiming to deliver pleasant aesthetic experiences. Still, Experiment 2 suggests that, once we fix a particular basic function for a role, the normative dimension represents that function as the ultimate aim of the commitment and not as a means to something else. In other words, the normative dimension represents commitments to basic functions as ends in themselves.

The plan for the rest of this paper is as follows. In section 2, we discuss some terminological issues, introduce the measure we use to study DCCs, and test its reliability in a preliminary study which we used to validate stimuli for our experiments. In sections 3-4, we present two experiments which test our account of the role of commitments in the normative dimension. In section 5,

we further develop our view, tie it to broader theories of concepts, and defend its explanatory adequacy by arguing that social role concepts represent commitments because that information is crucial to predict the future social roles and role-dependent behaviour of others.

2 Preliminary Study

Our study of DCCs focuses on social role terms such as *firefighter*, *bartender*, and *scientist*. Broadly speaking, these terms pick out professional social roles. We focus on this class for three reasons. First, we agree with Leslie (2015a) that social role concepts are the paradigmatic examples of DCCs.[2] Second, the result that most natural kinds are not, in their default readings, DCCs has been convincingly established by Knobe, Prasada, and Newman (2013). Third, there are important effects of gender on judgments involving DCCs, which we explore in follow-up work. To control for this, we excluded gender-based social roles such as *boy* and *mother*.

Turning to the measure of interest, Knobe et al. (2013) argue that one way to determine whether a term expresses a DCC is to consider its acceptability under *true*-modification. By '*true*-modification' we refer to sentences in which the predicate is modified by *true*, as in (3):

(3) Jack is a true scientist.

The basic idea is that, in this kind of linguistic construction, the *true*-modifier operates on the normative dimension. *Scientist* is a paradigmatic DCC, and (3) seems perfectly acceptable. For terms that arguably have either no normative dimension or no default value on it, *true*-modifications will seem less acceptable, as in (4)-(5):

(4) Jack is a true bus driver.

(5) Jack is a true second cousin.

As (4)-(5) illustrate, not all social role terms are equally acceptable under *true*-modification. Whether this means that only some social role terms are really DCCs—as Knobe et al. (2013) and Leslie (2015a) seem to think—is an issue that we shall discuss later. For now, let us simply call social role terms which receive low acceptability ratings in *true*-modifications 'weak DCCs' and those which receive high ratings 'strong DCCs'.

Although we distinguish weak and strong DCCs, it is also important to note that DCCs are sensitive to background information. As Leslie (2015a) argues, many weak DCCs can be turned, in certain conditions, into strong DCCs. More precisely: background information (e.g., the discourse context)

[2]Knobe et al. (2013) show that DCCs are also found in other conceptual domains, e.g., amongst some artifact kinds. Still, future studies are needed to determine the degree of uniformity in the content and structure of the normative dimension of DCCs which otherwise cut across different conceptual kinds.

can increase the acceptability ratings of *true*-modifications involving social role terms that, in other conditions, receive low ratings. This suggests that one way of investigating the content and structure of DCCs is to examine the interaction between background information and shifts in the acceptability of weak and strong DCCs in *true*-modifications. By tracing the parameters in background information which increase or decrease the acceptability of DCCs, we can test hypotheses about the content and structure of DCCs. We follow this strategy in Experiments 1 and 2, which we present in sections 3-4.

Given our strategy for investigating DCCs, we had to verify that acceptability ratings of *true*-modifications reliably order social role concepts such that some get lower ratings than others. To determine that, we replicated in this Preliminary study one of the basic results of Knobe et al. (2013), according to which different classes of social role concepts obtain reliably different ratings under *true*-modification. All subjects who participated in our experiments were recruited through Amazon's Mechanical Turk system for human intelligence tasks. Each participant was presented with a single question and was subsequently prompted to provide demographical data on age, gender, and native language. All participants were reimbursed for their participation. We excluded participants who were not native English speakers, were 17 years or younger, did not complete the survey, or had already filled out one of our surveys.

2.1 Method

257 participants were recruited for this preliminary study. Each participant received a single statement of the form *Jack is a true artist* without any context. Following Knobe et al.'s measure, participants were then asked to rate how natural/weird that sentence sounded on a 7-point Likert scale from '1' meaning 'sounds weird' to '7' meaning 'sounds natural'. We used all the professional role concepts that Knobe et al. classified as either DCCs or controls. Their DCCs included seven concepts: *artist*, *comedian*, *minister*, *musician*, *scientist*, *soldier*, and *teacher*. Their controls included eleven other professional concepts: *baker*, *bartender*, *caseworker*, *cashier*, *doorman*, *firefighter*, *mechanic*, *optician*, *tailor*, *waiter*, and *welder*. Each participant was randomly assigned to one of the 18 social role concepts.

2.2 Results

We first ran an analysis to determine whether the overall scaling of concepts according to their acceptability under *true*-modification was reliable. The analysis demonstrated high reliability: Spearman Brown's $\rho = 976$. We then investigated the differences between DCCs and social role controls. The mean rating for DCCs ($N = 111$) was 5.52 ($SE = 0.16$) and for social role controls ($N = 146$) was 4.71 ($SE = 0.15$). An independent t-test was performed, $t(255) = 3.74, p < .001$, indicating a highly significant difference between DCCs and social role controls. The average ratings for each of the concepts are presented in Figure 1.

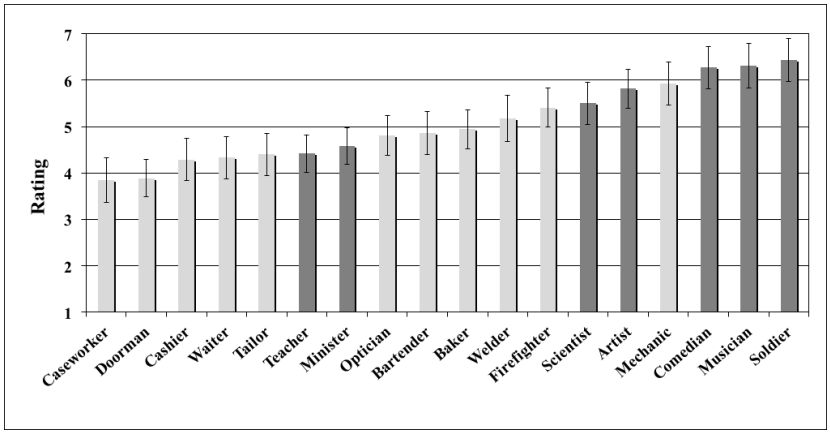

Figure 1. Acceptability-ratings for *true*-modifications of 18 social role concepts. The concepts were taken from Knobe et al. (2013). Dark grey columns identify concepts classified as DCCs in Knobe et al. (2013)'s study. Bars indicate the standard error around the mean.

2.3 Discussion

This Preliminary Study shows that social role concepts can be reliably ordered in terms of their acceptability ratings under *true*-modification. In addition, we replicated a key result of Knobe et al. (2013), according to which the mean ratings for DCCs such as *artist* is significantly higher than the mean ratings for social role controls such as *doorman*.[3] At the same time, if we consider individual ratings, we see a smooth transition between the acceptability ratings of the highest and lowest rated DCCs, as illustrated in Figure 1. This suggests that we should talk instead of strong and weak DCCs. In our subsequent experiments, we investigated the normative dimension of DCCs by examining how controlled manipulations of background information affect the acceptability ratings of *true*-modifications involving the social role terms presented in this Preliminary Study.

3 Experiment 1

As discussed in the Introduction, our view is that the normative dimension of a DCC does not directly represent its associated basic function; rather, it represents that function as the aim of a commitment. According to a natural competing view, the normative dimension represents the basic functions associated with social roles in a direct and unmediated way. Aside from some intuitive

[3]Although our study replicated the basic results of Knobe et al. (2013), some terms did change from weak to strong DCCs, and vice-versa, as is shown in Figure 1. This difference might be partly due to the fact that we only used a male protagonist in our stimuli, and we should briefly explain that choice. In the studies reported in Reuter and Del Pinal (2016), we found that most of the 18 concepts used by Knobe et al. (2013) are gender-associated (in an independent test) as either neutral or male. We then show that the gender of the protagonist has an important effect on the ratings of *true*-modifications. To control for that effect in the present studies, we used only male protagonists.

support, what is the evidence for our view? One might suspect that experiment 2 of Knobe et al. (2013) is relevant. However, their vignettes cannot be used to address this issue. To see why, consider their vignette for *scientist*:

> George is employed at Ameritech to run experimental studies and analyse the data. However, he actually has no interest at all in finding the correct answers to the questions he is studying. So although he goes through the motions, he does not actually care in any way about making a contribution to people's understanding of these issues.

This vignette does not assert that George is a good or skillful scientist; indeed, it implies quite the opposite. In particular, it seems to exclude the possibility that George fulfills the idealised basic function of a scientist ('he has no interest at all in finding correct answers...'). In general, Knobe et al.'s vignettes do not suggest that the relevant individuals are good or skillful at their job, hence they cannot be used to argue for the distinction between the functional and the normative dimensions of conceptual representation.

To test the hypothesis that what matters for the normative dimension is not whether someone is skilled or efficient, but rather whether someone is committed or devoted to the basic function, we presented participants with background information that asserted either the high-skill or the high-devotion of protagonists, and compared their effects on the acceptability ratings of *true*-modifications. For example, the high-skill condition highlights Jack's skill as an artist, while the high-devotion condition highlights Jack's devotion to being an artist. If the normative dimension represents unmediated basic functions, then information of someone's high skill should provide more direct evidence that the normative dimension is satisfied compared to information about someone's commitments. However, if our view is correct, the high-devotion condition should have a stronger positive impact on the acceptability of *true*-modifications than the high-skill condition. For on this view, information about someone's commitment to the function provides the more direct evidence of satisfaction of the normative dimension.

3.1 Methods

We used the 18 social role concepts tested in the Preliminary study. To examine how the high-skill and high-devotion conditions interact with the normative dimension, and how ratings in those conditions compare to the no background context (baseline) condition, we randomly assigned 899 participants to either one of the three conditions: no background context (N=303), high-skill (N=294), and high-devotion (N=302). The high-skill and high-devotion conditions, which included strong and weak DCCs, read as follows:

> (*High-skill*) Jack is an artist/doorman. He is very skilled and highly efficient at his job.

(*High-devotion*) Jack is an artist/doorman. He really likes his job and is completely devoted to it.

After being randomly assigned to one of the 3x18 conditions, participants were asked to rate whether the corresponding sentence of the form *Jack is a true artist/doorman* sounded weird (1), natural (7), or anything in between.

3.2 Results

To understand the analysis and results of Experiment 1, recall that, as shown in the Preliminary study, DCCs vary substantially in terms of their acceptability when there is no background information: weak DCCs get low ratings, strong DCCs get high ratings, and there is gradual progression between the extremes. For this reason, we have to examine differences in the effect of each background condition on ratings for weak and for strong DCCs. Our account predicts that the high-devotion condition should have a stronger positive effect than the high-skill condition on DCCs whose default rating in the no context condition has space for improvement. To test this, we used the average ratings of the 18 concepts in the Preliminary study to perform a median split, dividing the concepts into weak DCCs (*caseworker, doorman, cashier, waiter, tailor, teacher, minister, optician, bartender*) and strong DCCs (*baker, welder, firefighter, scientist, artist, mechanic, comedian, musician, soldier*). As reported in the Preliminary Study, the scaling of the concepts in terms of acceptability ratings under *true*-modification was highly reliable. Indeed, only a single concept that was classified as a weak DCC in the Preliminary Study received a rating that placed it amongst the stronger DCCs in the no context condition of Experiment 1. The average ratings for weak and strong DCCs in all three conditions of Experiment 1 are shown in Figure 2, and the average values for all individual concepts in each of the three conditions are presented in Table 1 of the Appendix. We conducted a nested ANOVA with participants' ratings as the dependent measure and the independent factors *Context* (no context, high-skill, high-devotion), *Concept-type* (strong DCCs, weak DCCs), as well as *Concept* (all 18 social role concepts, random) nested within *Concept-type*. A highly significant interaction occurred between *Context* and *Concept-type*, $F(2, 845) = 9.71, p < 0.001$. Crucially, for weak DCCs, the high-devotion condition significantly increased their average rating from 4.30 (in the no context condition) to 5.09, whereas the high-skill condition non-significantly increased their average ratings to 4.50. This result confirms our prediction. For strong DCCs, which are close to ceiling in the no context condition, the high-devotion condition had no effect, and the high-skill condition *decreased* their ratings from 5.74 to 4.97.

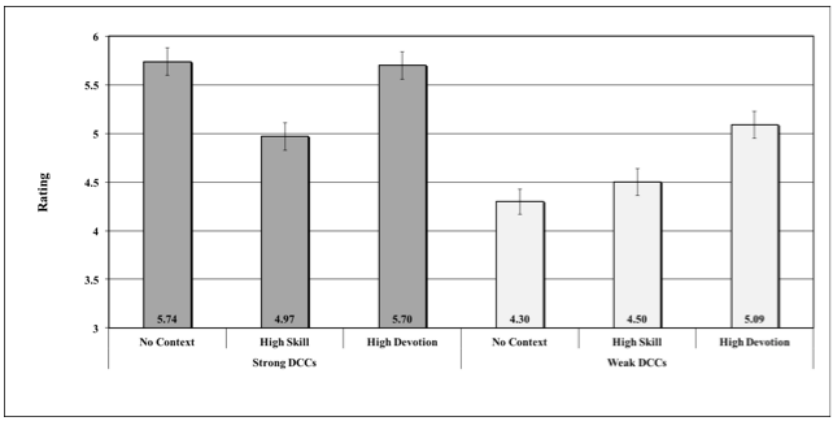

Figure 2. Acceptability-ratings for strong and weak DCCs in the no-context, high-skill, and high-devotion conditions. Bars indicate the standard error around the mean.

Our account also predicts a main effect of *Context* such that the high-devotion condition should have a stronger effect on the average acceptability of *true*-modifications compared to the high-skill condition. The analysis revealed a significant main effects for *Context*, $F(2, 845) = 16.71, p < 0.001$. Averaging over the 18 social role concepts, the high-devotion condition raised the mean value from $5.01(SE = 0.10)$ to $5.39(SE = 0.09)$, whereas the high-skill condition dropped the value to $4.74(SE = 0.10)$. Posthoc tests revealed a significant difference between the high-devotion and the high-skill conditions ($p < 0.001$), as well as between the high-devotion and the no-context conditions ($p = 0.017$), whereas no significant effect was found between the high-skill and the no-context conditions ($p = 0.124$). A significant main effect was found for *Concept-type*, $F(1, 845) = 27.21, p < 0.001$, which is expected given the basis for the classification of concepts into weak and strong DCCs. The analysis also revealed a significant main effect for the nested factor *Concept*, $F(16, 845) = 2.86, p = 0.005$, indicating that there is substantial variation in the *true*-ratings amongst the individual concepts classified as weak and strong DCCs. No interaction occurred between *Context* and the nested factor *Concept*, $F(32, 845) = 0.71, p = 0.885$.

As shown in Figure 3, we created two scatterplots with data from the no-context condition on the x-axis, and the results from the high-devotion and high-skill conditions on the y-axis, and fitted the data points using linear models. The fitted lines demonstrate that progressively higher ratings in the no-context condition correspond with progressively higher ratings in the high-skill and high-devotion conditions. This indicates that participants processed the information in each condition as intended: specifically, they do not just process the information provided by the background context, but also that provided by the specific social roles. Importantly, Figure 3 also shows that, for almost every concept, the average ratings in the high-devotion condition are greater than in the high-skill condition. This reveals that the main effect of *Context* reported in the ANOVA is not driven by a special subset of the concepts, since there is high consistency in the ratings for the 18 concepts across both context conditions.

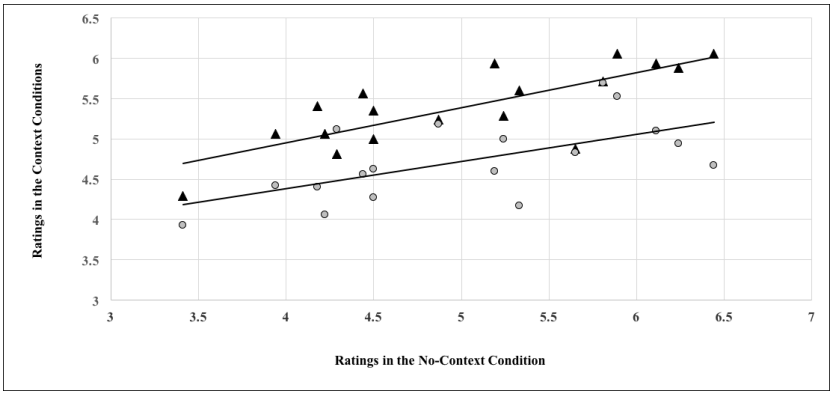

Figure 3. Scatterplot for all concepts. Triangles indicate the rating of each concept in the high-devotion condition and circles indicate the rating in the high-skill condition.

3.3 Discussion

Experiment 1 examines how background information regarding the level of skill or commitment of protagonists affects the ratings of *true*-modifications involving DCCs. According to our view, which says that the normative dimension represents basic functions as the objects of commitments, information of the protagonist's high commitment or devotion should have a significantly stronger positive effect on the ratings. This effect should be most clearly observed in the case of weak DCCs, since these concepts obtain relatively weak ratings when there is no background information.

Confirming this prediction, the most general result of Experiment 1 is that, compared to a control condition with no background information, information of high devotion significantly increased the overall acceptability of *true*-modifications. In contrast, information of high skill and efficiency in fulfilling the roles did *not* increase the overall acceptability of the corresponding *true*-modifications. This pattern of results directly supports the hypothesis that the normative dimension of DCCs represents commitments to fulfill the relevant basic functions. This view is further reinforced by the fine-grained analysis which looked into the effects of the high devotion vs. the high skill conditions on the ratings of *true*-modifications for strong vs. weak DCCs. Strong DCCs such as *scientist* and *artist* already obtain high acceptability ratings under *true*-modification in the no-context condition, as shown in the Preliminary Study and further confirmed in the no context condition of Experiment 1. Since the ratings of strong DCCs are at or close to ceiling, there is little scope to observe the predicted positive effect of information of high-devotion. In contrast, weak DCCs such as *doorman* and *cashier* obtained relatively low ratings in the no context condition, hence can be used to observe the effects on the ratings of additional information about the high devotion or skill of protagonists. As predicted by our hypothesis, the high devotion condition had a significant positive effect on the acceptability ratings of weak DCCs, whereas the high skill condition did not significantly increase their acceptability.

As we just said, strong DCCs obtain high default ratings in the no context condition, and for this reason are not directly useful in terms of testing the competing accounts of the normative dimension. It is unsurprising, then, that the high devotion condition had no effect on the ratings of strong DCCs. However, the high-skill condition had a *negative* effect on the acceptability ratings of strong DCCs, which resulted in an interaction between *Concept-type* and *Context*. This negative effect is not predicted by the accounts of the normative dimension considered here, and seems to require an independent explanation. One possibility—compatible with our account—is that the drop is due to a pragmatic effect. Specifically, saying that Jack is a skilled artist can be taken as suggesting that he is *at most* only a skilled artist, hence that he lacks something which is relevant to his being a true artist.[4]

To sum up, Experiment 1 shows that, for DCCs whose default ratings have scope for improvement, including weak DCCs, information of the protagonist's high-devotion has a significant positive effect on the acceptability of *true*-modifications. In contrast, information of the protagonist's high-skill has no significant positive effect on the acceptability of *true*-modifications. This pattern of results support the hypothesis that the normative dimension of DCCs represents the commitments to fulfill the basic function of the corresponding social roles, and not just the unmediated basic functions by themselves.

4 Experiment 2

Experiment 1 supports the view that the normative dimension of DCCs involves the representation of commitments. Now, commitments in general, and their representations in particular, have structure, i.e., they are commitments to something. We argued that the normative dimension of DCCs represents commitments as aiming to fulfil the basic function of the corresponding social role. However, Experiment 1 does not tell us much about the aims of these commitments. For example, maybe it matters for being a true artist that Jack tries to produce works of aesthetic value, but it is irrelevant whether he does so because that is the ultimate aim of his commitments qua artist, or because he believes that is a means to become rich or famous. The question, then, is whether DCCs represent commitments to their basic functions as ends in themselves, as we hold, or whether they can also be commitments to other ultimate aims, such as fulfilling the role as a means to get money or fame.

[4] Another possibility is that, when presented with a high-skill condition, participants ignore the content of the social role term, and focus on the relation between the context and the modifier *true*. As a result, we observe, relative to the no context condition, a drop in the ratings for strong DCCs and an increase for weak DCCs. The problem with this suggestion is that there is no principled reason why it shouldn't also apply to the high-devotion condition. The ratings for the high-skill condition should be roughly the same across weak and strong DCCs, and similarly the ratings for the high-devotion condition should be roughly the same across weak and strong DCCs. However, this is not what we observe: the ratings of DCCs in the high-skill and high-devotion conditions are significantly positively correlated with the ratings in the no-context condition (see Figure 3).

In addition, we have presented our view as if each social role term is associated with a fixed basic function. This is a substantial simplification. Indeed, any plausible view of DCCs must be compatible with the well-known fact that the concepts expressed by particular words can vary depending on the discourse context, among other things (Roth and Shoben, 1983; Barsalou, 1987; Rosch, 2011). In particular, we expect that the basic functions associated with social role terms can, to some extent, vary in response to background information. We might have a default concept associated with *artist*, but we often use this term to talk about more specific or even idiosyncratic subclasses: e.g., it can express a concept closer to that of a composer, a painter, or a performer, and its associated function can vary accordingly.

The relative context-sensitivity of social role terms provides us with a way to test our hypothesis regarding the way in which the normative dimension of DCCs represents commitments. If our view is correct, once a function is associated with a DCC on a particular occasion, the commitment to that function is represented as an end in itself. The alternative view is that simply having commitments matter, regardless of what their ultimate aim is—so long as the role and its function is a reasonable way to achieve that ultimate aim. To test these hypotheses, we provided participants with vignettes that vary two parameters: the basic function that Jack performs as a member of certain social roles, and the ultimate aim of Jack's commitment qua member of that role. In some vignettes, the function is the ultimate aim of the commitment, and in others the function is only a means to some other ultimate aim, as in the example of the artist who creates art to get money or fame. If our hypothesis is correct, cases in which the function is the ultimate aim of the commitment should result in higher acceptability ratings of the corresponding *true*-modifications compared to cases in which the function is only a means to some other end.

4.1 Methods

811 participants were recruited for Experiment 2. We created vignettes in which we varied both the basic *Function* of the protagonist's social role (*Default function*, *Preference*, *Output*), and the ultimate *Aim* of his role (*End in itself*, *Money*, *Fame*). To illustrate, here is the example for *artist*:

> Jack is an artist. He is talented and very skilled. Jack tries to produce
>
> a. works of art that inspire other people. [*Default function*]
> b. works of art that he likes. [*Preference*]
> c. as many works of art as possible. [*Output*]
>
> He aims to do so because he believes that
>
> (i) this is the purpose of an artist. [*End in itself*]
> (ii) those works of art will earn him a lot of money. [*Money*]
> (iii) those works of art will make him famous. [*Fame*]

Participants were presented with a vignette obtained by matching one of the *Functions* in (a)-(c) with one of the *Aims* in (i)-(iii), and were then asked

to rate whether the sentence *Jack is a true artist* sounded weird (1), natural (7), or anything in between. Note that the vignettes that involved the aim *End in itself*—i.e., a-(i), b-(i), c-(i)—were the conditions in which the ultimate aim of the commitment is the basic function. The other vignettes were the conditions in which the ultimate aim is something else, namely, *Money* or *Fame*. To ensure the comparability of the responses, our vignettes *only* varied the parameters *Function* and *Aim*. These vignettes could be coherently applied to six professional roles: *artist*, *baker*, *comedian*, *musician*, *scientist*, *tailor*, all of which share the feature of producing identifiable social goods, and of being reasonable means to get money or fame. Each participant was randomly presented with one of the 3 x 3 x 6 conditions. The full set of vignettes is presented in the Appendix.

4.2 Results

Figure 3 below depicts the mean values for the nine conditions averaged over all six DCCs. We performed a 3 x 3 x 6 ANOVA with *Aim* (end-in-itself, money, fame), *Function* (default, preference, output), and *Concept* (*artist*, *baker*, *comedian*, *musician*, *scientist*, *tailor*) as independent factors, and participant's rating as the dependent measure. The data exhibited significant main effects for *Aim*, $F(2, 757) = 23.57, p < 0.001$, for *Function*, $F(2, 757) = 4.16, p = 0.016$, as well as for *Concept*, $F(5, 757) = 4.49, p < 0.001$. The overall average ratings for each of the three *Aims* were $5.39(SE = 0.11)$ for end-in-itself, $4.59(SE = 0.11)$ for money, and $4.58(SE = 0.10)$ for fame. Posthoc tests for *Aim* revealed highly significant ($p < 0.001$) differences between end-in-itself and money, as well as between end-in-itself and fame, but not between money and fame ($p = 1.000$). Posthoc tests for *Function* showed a significant difference between default function and output ($p = 0.008$), but not between default function and preference ($p = 0.135$), or between output and preference ($p = 0.937$). No significant interaction occurred between *Aim* and *Function*, $F(4, 757) = 0.66, p = 0.616$. To investigate the internal consistency across concepts, we determined possible interactions between *Aim* and*Concept*, and *Function* and *Concept*. While no significant interaction was found between *Aim* and *Concept*, $F(10, 757) = 1.58, p = 0.108$, a significant interaction was observed between *Function* and *Concept*, $F(10, 757) = 2.87, p = 0.002$.

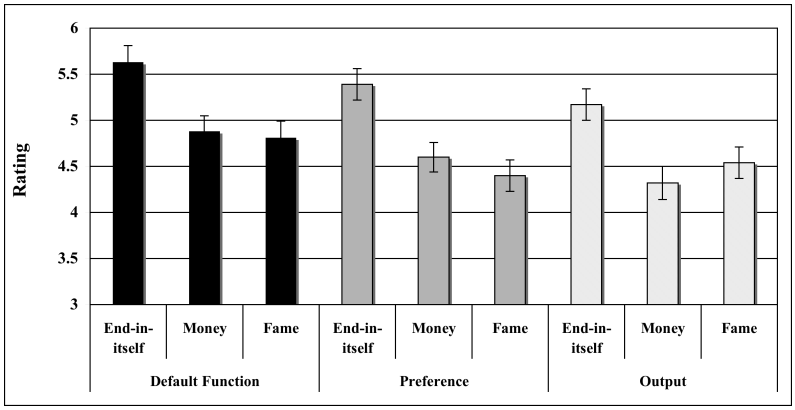

Figure 3. Comparison of average ratings for three different basic functions—default function, preference, output—when that function is described as being the ultimate aim of the commitment (end in itself), or as being a means to something else (money, fame). Bars indicate the standard error around the mean.

4.3 Discussion

Experiment 2 explores the relation between the basic function of social role terms and the structure of commitments as represented in the normative dimension. In our view, the normative dimension of DCCs has a fixed, 'end in itself' structure: the basic function of the DCC is the ultimate aim of the commitment. According to the competing view, the normative dimension can have a means-end structure: the aims of the commitments allow for the basic function of the DCC to be the means to some other end. Experiment 2 obtained three main results which, taken together, strongly support our view.

The first result is that participants are relatively flexible in the type of basic function that they can associate with particular social roles. The main function associated with an *artist* can be to produce works that the artist thinks others will find inspiring, or simply works that he likes. Now, we are not suggesting that anything goes. For example, assigning an artist the basic function of buying art would be bizarre and likely unacceptable to most participants. Indeed, our results show that, for most concepts, the basic default function had a stronger positive effect on the ratings compared to both output or preference. However, the acceptability pattern for all three functions varies considerably across the six social role concepts. For example, while output (producing lots of theories) had the lowest rating for *true*-scientist, output (efficient tailoring) had the highest rating for *true*-tailor. Taken on its own, this result is consistent with both views under discussion.

The second result is that, given any of the three kinds of basic functions for a social role, participants are more willing to accept the corresponding *true*-modification whenever the ultimate aim of the commitment is simply to fulfill that social role's function. In particular, if Jack's aim in performing the role of an artist is, ultimately, to acquire money or fame, then the ratings of the corresponding *true*-modification drop significantly. The general result is clearly

illustrated in Figure 3. Independently of which type of basic function is selected (default, output, or preference), the *Aim* that always gets the highest ratings is end-in-itself, i.e., the condition in which the ultimate aim of the commitment is to fulfill the role's function. This result was significant and consistent across all six social role concepts, hence clearly supports our claim regarding the 'end in itself' structure of the normative dimension. In addition, it is in tension with the alternative view which allows commitments to basic functions to be represented as means to some other end, such as making money or seeking fame.

The third key result is that there is no interaction between the type of basic *Function* and *Aim*. This suggests that, regardless of the type of function which we associate with a social role (default, output, or preference), the structure of the commitment, as represented in the normative dimension, is fixed as an end in itself commitment to the role's basic function. On the competing view, the ultimate aim of the commitment could be to fulfill the basic function as a means to make money or pursue fame. We saw that this view is not supported by the second result. Still, we can imagine a variant that says that the structure of commitments can vary depending on the type of function of the DCC. For example, if the basic function of *artist* is to inspire people, then the ultimate aim of the commitment is to that function, but if the function is productivity, then the ultimate aim could reasonably be to make money or seek fame. Crucially, however, this view does not cohere with the result that no interaction was found between type of function and the way in which commitments are represented. The representation of commitments seems to have a fixed, end in itself structure: what matters for satisfying the normative dimension of a social role concept is an ultimate commitment to the role's function.

5 General Discussion

Experiments 1 and 2 provide us with key insights into the content and structure of the normative dimension of the DCCs expressed by social role terms. We have argued that the normative dimension represents commitments to the basic function associated with the relevant social roles. We have also argued that such commitments are represented in a relatively constrained way, namely, as ends in themselves. To further develop and refine this view, we now briefly address three important general questions. (i) How should we integrate our account of DCCs into broader accounts of the structure of lexical concepts? (ii) How do the normative and other dimensions of DCCs interact with linguistic modifiers and non-linguistic background knowledge? (iii) Is our theory of the normative dimension explanatorily adequate? In other words, why should social role concepts include a normative dimension that represents commitments to basic functions?

5.1 Conceptual structure of DCCs

A key task in the study of DCCs is to determine how best to fit them into broader theories of lexical concepts (Knobe et al., 2013). One suggestion is that DCCs are a subclass of essentialist concepts alongside natural kinds. According to essentialists, we posses 'folk theories' in which essences are not only what define certain categories but are also causally responsible for their concrete or typical features (Keil, 1989; Gelman and Wellman, 1991).[5] Although the relation between the normative dimension and the concrete features of DCCs is one of 'realization' rather than 'causation', one could adopt an analogous view according to which we also have 'normative folk theories'. The concrete features of DCCs are then unified as being those features that, according to our normative theories, typically manifest the corresponding commitments represented in their normative dimension. On this view, although we often use concrete features to categorize entities as DCCs, we understand that the deeper, independent, and final criterion of membership is the abstract value.

Although we share the intuition that the normative dimension is a central dimension of DCCs, we doubt that it is akin to an essence. Experiment 2 of Knobe et al.'s study shows that there is a key difference between the role of the essence in categorization with natural kinds and the role of the normative dimension in categorization with DCCs. As illustrated by (1-a) and (1-b) above, someone can fail to satisfy the normative dimension but still fall under a DCC; but the analogous pattern does not hold for natural kinds, the paradigmatic examples of essentialist concepts. Imagine an animal that looks and hunts like a tiger, but gives birth to pumas and has puma DNA. Consider (6-a) and (6-b):

(6) a. There is a sense in which that animal is clearly a tiger.
b. However, if you think about what it really means to be a tiger, you would have to say that it is not really a tiger.

To most people, (6-b) will seem fine, but (6-a) will seem unacceptable. Experiment 2 of Knobe et al. (2013) confirms this intuition. Essentialist accounts have a good explanation for this response pattern, since they hold that the real criterion for membership in a natural kind is satisfaction of the essence. In this case, there is direct evidence that the animal does not have the tiger essence; so despite being superficially like one, there is no sense in which it is a tiger.[6]

[5] In this discussion of the 'essentialist' theory of concepts, we focus on a traditional version of the theory. In an instructive discussion, Strevens (2000) distinguishes between three versions of essentialism: pure, internal and statistical essentialism. Pure and internal essentialists assume that the core is what defines categories in the sense that satisfaction of the core is necessary and sufficient for membership. Strevens argues that traditional essentialists such as Keil (1989) and Gelman and Wellman (1991) are pure or internal essentialist. This is the version of the essentialism we will explicitly reject when extended to DCCs.

[6] For interesting empirical investigations of differences in the features, dimensions, and the way in which they are integrated in categorization under natural kinds compared to artifact kinds, see Hampton et al. (2007) and Hampton et al. (2009). For our purposes, we are only pointing out the observed contrast between DCCs and natural kinds in cases analogous to (6-a)-(6-b). We are not endorsing a traditional essentialist theory for natural kinds, which the studies just mentioned show would require some substantial reformulation.

The normative dimension of DCCs, then, behaves more like a central but defeasible dimension than like an essence in a natural kind. To say that a dimension or feature d is central in a concept C is to say that other dimensions and features of C depend on d more than it depends on them. However, to say that, despite its centrality, d is not an essence is to say that, if a sufficient number of other features are satisfied, something can be a C despite not falling under d. One way to capture this suggestion is to adopt a version of prototype theory that organizes features not only into sets of weighted dimensions, but also encodes the dependency relations amongst those dimensions (Sloman et al., 1998). Hampton (2006) calls these structures 'theory-based prototypes'. To be clear, neither Experiments 1-2, nor any of the experiments presented by Knobe et al. (2013), provide the data we need to determine, with precision, the relative centrality of the normative dimension in the conceptual structure of DCCs. We focus on this issue in future work.[7]

5.2 DCCs in context

To sketch how DCCs interact with linguistic modifiers, consider first some general properties of their conceptual structure. Take the lexical entry for *scientist*. Based on what we said so far, this entry includes, at least, the following dimensions: PERCEPTUAL (how scientists look), FUNCTIONAL (function$_{d/c}$ of scientists), and NORMATIVE (commitment to fulfill the function$_{d/c}$ of scientists). The subscript 'd/c' in 'function' captures part of the structural constraint which we established in Experiment 2. Terms like *scientist* have a *default* basic function$_d$, which can change to a *contextually* relevant basic function$_c$. When that happens, the aim of the commitment changes from function$_d$ to function$_c$, but, crucially, it continues to be represented as an end in itself. The DCCs which represent social roles likely include other dimensions (Pustejovsky, 1995; Del Pinal, 2015), which for simplicity we ignore here.[8]

As we mentioned above, we assume that conceptual representations can encode the relative weights of, and the dependency relations between, features and

[7] One could hold that social role terms might be ambiguous and express a descriptive social role concept in one sense and a purely normative concept in another. We empirically address this worry in our studies of centrality: if it turns out that the normative dimension is central to social role terms (as indeed our initial results suggest), this would support the view that it is a dimension in a unified conceptual structure used to represent social roles. For theoretical reasons to hold that these dimensions are most useful if unified, see §5.3 below.

[8] Some philosophers would argue that those features or dimensions which are not necessary for determining the reference or extension of concepts or terms should not, strictly speaking, be taken as part of their conceptual structure. From this perspective, most empirical studies of conceptual structure, such as those found in Knobe et al. (2013) and the studies presented above, are really studies about our 'conceptions' of classes (cf. Burge, 1993). Our account should still be of interest to those philosophers, at least if they also agree that it is important to understand how we 'conceive' of social roles. Still, it is crucial to note that, as we will explain below, the compositional behaviour of social roles does seem to depend on linguistic operations having access to such 'conceptions'. If this is correct, conceptions will do some of the job that concepts traditionally are thought to do. For a detailed defence of this view, see Del Pinal (2015).

dimensions (Sloman et al., 1998; Hampton, 2006; Hampton et al., 2009). Focusing on the weights and membership thresholds of DCCs, we have seen that, for most DCCs, someone can count as a member by satisfying the NORMATIVE, but also by satisfying the FUNCTIONAL.[9] Furthermore, although either dimension might be sufficient, neither one seems to be necessary for membership. At the same time, it seems clear that someone cannot be, say, a scientist if s/he satisfies the PERCEPTUAL, but neither the NORMATIVE nor the FUNCTIONAL.

Now, conceptual structures can be modified by linguistic operators. In the framework we adopted, a natural way of modelling these effects is as follows. Modifiers such as *efficient* and *true* take conceptual structures as inputs and change them in certain systematic ways (cf. Smith et al. 1988; Hampton 1995). An *efficient scientist* is efficient at producing experiments and theories. A *true scientist* is someone who is committed to the aims of science. In general, the effect of each modifier is to change the membership threshold, centrality relations and weights of the dimensions of the modified noun. For example, *efficient* increases the relative weight of the FUNCTIONAL so that it becomes necessary to pass the threshold, and *true* increases the relative weight of the NORMATIVE so that it becomes necessary to pass the threshold.

Other modifiers can be treated in a similar way. One interesting modifier, extensively discussed in Leslie (2015a), is *real* as used in expressions such as *John and Mark are painters; but only John is a real artist.* Leslie points out that *real* and *true* both operate on the normative dimension. There might, however, be a subtle difference. *True* increases the relative weight of the NORMATIVE, so there are cases in which someone can be an amateur scientist but still be a true scientist, if his commitments are in the right place. In the case of *real*, both dimensions seem to be crucial, i.e., to be a *real scientist* both the NORMATIVE and the FUNCTIONAL have to be satisfied. Finally, the use of linguistic modifiers is not necessary to perform these changes in the structure of DCCs; as Leslie emphasises in her discussion of slurs such as *Hillary Clinton is the only man in the Obama Administration*, non-linguistic information can also affect the structure of social role concepts in analogous ways.

5.3 Why do social role concepts have a normative dimension?

We argued that the concepts expressed by many social role terms have a normative dimension which represents commitments to basic functions. Insofar as our account coheres with the available data, we can say that it is descriptively adequate. At the same time, one might inquire as to the explanatory adequacy of this account. Why should there be a normative dimension, so conceived, in the conceptual structure of DCCs at all? In particular, why should social roles, but

[9] To illustrate, recall the initial example involving John, the biological and successful father, and Mark, the non-biological and committed father. Those examples suggest that most people are willing to call someone a father, as long as he satisfies at least one of the FUNCTIONAL or the NORMATIVE.

not other types of concepts, have a dimension involving commitments? To meet the demands of explanatory adequacy, we must address these sorts of questions.

The structure and information encoded in concepts, most cognitive scientists would agree, is largely determined by two factors (Rosch, 1999). (i) Our conceptual system aims to provide maximum information about the environment with the least cognitive effort, and in particular to predict as many properties as possible from knowing any one property. (ii) We perceive properties in the world as having high-correlational structure. If we apply these principles to the human social world, we can begin to see why, in that domain, representing commitments turns out to be useful.

Suppose you want to predict the behaviour of objects y and z. If you learn that y is a gold ring and z is a bee, you can make many predictions: e.g., that you can grab y but should leave z alone, and that neither one is as tasty as an apple. The behaviour of the bee is more complex, and requires some form of means-ends reasoning. For everyday purposes, the set of basic goals which we use to interact with a random bee derive from their kind and the local environment. For more refined purposes, we can subdivide into worker bees, queen bees, etc., which will determine a more refined set of functions. Still, even in these cases there is no use in representing bees in terms of their 'commitments'.

However, once we add highly flexible social entities such as humans unto the perceived environment, we add an important layer of complexity. Human behaviour is radically unpredictable from the sorts of general biological facts we use to understand other animals. To predict much of human behaviour in any useful way, one needs to know more than just their biological properties and relevant environmental facts. In particular, one needs to know their social roles and corresponding functions. Furthermore, one key property of many social roles is that, in certain conditions, they are voluntary. Since a random Jack can choose to be an artist or scientist, and since Jack can also choose to change roles, it is particularly useful to categorize Jack in terms of his commitments.

Given these basic properties of human behaviour, information about the relevant commitments is fundamental to make useful categorizations and predictions. In particular, it is useful to know not only whether Jack is good or bad at something, but also whether he is committed to it. Suppose you know that Jack is a really good pianist. Will he become a professional pianist? Hard to say. Suppose that you know, in addition, that Jack is committed to being a pianist. Then predicting his future behaviour is easier. Suppose that John is not, at the moment, a good pianist. Will he become a professional pianist? Probably not, but still hard to say. Suppose you then find out that he is committed to being a pianist. You can predict, with some confidence, that he will keep on playing, despite currently being quite bad at it.

Almost everyone is, at some point in their development, quite bad at performing the functions of their future social roles. However, whether they will eventually perform those functions, whether they will become good at them, and if good, whether they will continue those activities, depends to a large extent on their commitments. So the information encoded in the FUNCTIONAL and in the NORMATIVE allows us to make predictions about how someone will behave in

certain conditions, and there are contexts in which one or the other dimension is more useful. Hence, despite the high-level talk of 'abstract values' and 'normative dimensions', there is nothing mysterious about why, given the basic function of categorisation and the kind of flexibility inherent to human behaviour, many of our social role concepts should include a dimension that encodes the relevant commitments. Without access to this dimension, our predictions regarding the behavior of others would be substantially impoverished.

This explanation of the usefulness of representing commitments in our concepts of social roles naturally accommodates the finding that there are weak and strong DCCs (see the Preliminary Study above). In our view, DCCs capture the basic structure of the perceived human social world, including sets of co-occurring features. Now, most people believe that gaining and/or keeping membership in some social roles requires a greater degree of commitment and effort than gaining and/or keeping membership in others. Our social role concepts seem to encode these differences. If we compare paradigmatic weak DCCs (e.g., *doorman*, *cashier*) with strong DCCs (e.g., *scientist*, *musician*), we can see that they likely reflect some common cultural evaluations along these lines. Crucially, those default assumptions can be reconfigured when the level of commitment of particular members is made explicit (see Experiment 1).

This explanation also sheds light on why the NORMATIVE seems to represent commitments to basic functions as ends in themselves, and not as means to other ends (see Experiment 2). Representing commitments helps us predict whether, say, Jack will become a pianist. To efficiently achieve that, the concept of a pianist directs our attention to information about the extent to which Jack is committed to being a pianist for its own sake, rather than as a means to make money or become famous. If Jack is already a pianist, assertions that Jack is a *true pianist* provide us with information about Jack's ultimate commitment to the basic function of a pianist. This information is crucial to predict Jack's likelihood to continue in that role, especially when certain externalities change its effectiveness as a means to other ends. The underlying bet of our conceptual system seems to be that, overall, an important predictor of someone's eventual or continual membership in a social role is the level of her/his commitment to that role's basic function as an end in itself.

To be clear, our claim is not that information about, say, Jack's commitment to a role is more important, in some absolute sense, than information about his skill at that role, for prediction and counterfactual thought. Our claim is rather that both are important, but their respective usefulness is crucially different. This is easy to see in cases of counterfactual reasoning. Suppose Jack is a very skilled doctor, and imagine that he is suddenly offered another job, which pays much more. What will Jack do? In this kind of scenario, information about Jack's level of commitment to being a doctor becomes highly relevant. Clearly, we often have to predict the decisions of others in structurally similar scenarios. Overall, these considerations suggest that our account of the content and structure of the normative dimension has the resources to meet the demands of explanatory adequacy.

5.4 Future directions

The empirical study of social role concepts and DCCs is in its infancy. To conclude, we briefly three promising areas for future research. First, it is reasonable to think that there are limits to what we intuitively accept as reasonable ends that beings like us can be genuinely committed to. Indeed, there are many social roles, different from the ones we examined here, whose basic function carries a negative connotation. In many of these cases, modification by *true* feels strange or inappropriate: e.g., consider *true prisoner* and *true money launderer*. An intriguing question is to explore the ways in which moral values and other normative considerations affect what we consider possible ends that individuals could be freely committed to. Second, when considering the general structure of social role concepts, how central, compared to other dimensions, is the NORMATIVE? We have argued that the NORMATIVE is unlikely to be as central for social roles as the essence is for natural kinds. Furthermore, the answer to this question could depend on the subclass of social roles being considered. Still, it is important to determine the relative centrality of the FUNCTIONAL and the NORMATIVE for each natural subclass of social role concepts. The third and final issue we want to mention is the following. Assuming we are correct about the role of commitments in our representations of social roles, what are the implications for more general issues in social cognition, including the study of gender and other biases? We show in Reuter and Del Pinal (2016) that varying the gender of the protagonist affects the ratings of *true*-modifications. Examining these sorts of effects might shed light on implicit assumptions regarding gender, commitments, and social roles. At this point, we hope to have shown, in the spirit of Knobe et al. (2013) and Leslie (2015a), that the detailed empirical and theoretical study of social role concepts promises to provide key insights into the nature of social cognition.

Appendix

A.1 Tables

DCC	No Context	High Devotion	Difference	High Skill	Difference
Cashier	3.41	4.29	0.88	3.93	0.52
Minister	3.94	5.06	1.12	4.42	0.48
Optician	4.18	5.41	1.23	4.40	0.22
Tailor	4.22	5.06	0.84	4.06	-0.16
Bartender	4.29	4.81	0.52	5.12	0.83
Caseworker	4.44	5.56	1.12	4.56	0.12
Waiter	4.50	5.00	0.50	4.63	0.13
Doorman	4.58	5.35	0.77	4.27	-0.31
Welder	4.87	5.24	0.37	5.18	0.31
Mechanic	5.19	5.94	0.75	4.60	-0.59
Teacher	5.24	5.29	0.05	5.00	-0.24
Baker	5.33	5.60	0.27	4.17	-1.16
Firefighter	5.65	4.88	-0.77	4.83	-0.82
Comedian	5.81	5.71	-0.10	5.69	-0.12
Soldier	5.89	6.06	0.17	5.53	-0.36
Artist	6.11	5.94	-0.17	5.10	-1.01
Scientist	6.24	5.88	-0.36	4.94	-1.30
Musician	6.44	6.06	-0.38	4.67	-1.77

Table 1. Average ratings for all 18 concepts in the three conditions used in Experiment 1, including difference scores between the high-devotion/high-skill condition and the no background

condition.

A.2 Vignettes used in Experiment 2

In Experiment 2, we obtained vignettes for each term by matching each of (a), (b), and (c) with each of (i), (ii), and (iii). Participants were randomly assigned to one of the 3 x 3 x 6 conditions. The full scheme for the vignettes are the following:

Artist

Jack is an artist. He is talented and very skilled. Jack tries to produce

a. works of art that inspire other people.
b. works of art that he likes.
c. as many works of art as possible.

He aims to do so because he believes that

(i) this is the purpose of an artist.
(ii) those works of art will earn him a lot of money.
(iii) those works of art will make him famous.

Baker

Jack is a baker. He is talented and very skilled. Jack tries to bake

a. cakes that people find delicious.
b. cakes that he likes.
c. as many cakes as possible.

He aims to do so because he believes that

(i) this is the purpose of a baker.
(ii) those cakes will earn him a lot of money.
(iii) those cakes will make him famous.

Comedian

Jack is a comedian. He is talented and very skilled. Jack tries to produce

a. jokes that make people laugh.
b. jokes that he likes.
c. as many jokes as possible.

He aims to do so because he believes that

(i) this is the purpose of a comedian.
(ii) those jokes will earn him a lot of money.
(iii) those jokes will make him famous.

Musician

Jack is a musician. He is talented and very skilled. Jack tries to compose

a. songs that captivate other people.
b. songs that he likes.
c. as many songs as possible.

He aims to do so because he believes that

(i) this is the purpose of a musician.
(ii) those songs will earn him a lot of money.
(iii) those songs will make him famous.

Scientist

Jack is a scientist. He is talented and very skilled. Jack tries to construct

a. theories that expand human knowledge.
b. theories that he likes.
c. as many theories as possible.

He aims to do so because he believes that

(i) this is the purpose of scientist.
(ii) those theories will earn him a lot of money.
(iii) those theories will make him famous.

Tailor

Jack is an tailor. He is talented and very skilled. Jack tries to tailor

a. costumes that suit his clients.
b. costumes that he likes.
c. as many costumes as possible.

He aims to do so because he believes that

(i) this is the purpose of a tailor.
(ii) those costumes will earn him a lot of money.
(iii) those costumes will make him famous.

References

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization*, pp. 101–140. Cambridge, UK: Cambridge University Press.

Burge, T. (1993). Concepts, definitions and meaning. *Metaphilosophy 24*(4), 309–325.

Del Pinal, G. (2015). Dual content semantics, privative adjectives and dynamic compositionality. *Semantics and Pragmatics 8*(3), 1–53.

Gelman, S. A. and H. W. Wellman (1991). Insides and essences: early understanding of the non-obvious. *Cognition 38*, 213–244.

Hampton, J. A. (1995). Testing the prototype theory of concepts. *Journal of Memory and Language 34*(5), 686–708.

Hampton, J. A. (2006). Concepts as prototypes. *The Psychology of learning and motivation: Advances in research and theory 46*, 79–113.

Hampton, J. A., Z. Estes, and S. Simmons (2007). Metamorphosis: Essense, appearance and behavior in categorization of natural kinds. *Memory & Cognition 35*, 1785–1800.

Hampton, J. A., G. Storms, C. L. Simmons, and D. Heusen (2009). Feature integration in natural language concepts. *Memory & Cognition 37*(8), 1150–1163.

Keil, F. C. (1989). *Concepts, Kinds and Cognitive Development*. Cambridge, MA: The MIT Press.

Knobe, J., S. Prasada, and G. E. Newman (2013). Dual character concepts and the normative dimension of conceptual representation. *Cognition 127*, 242–257.

Leslie, S.-J. (2015a). "Hilllary Clinton is the only man in the Obama administration": Dual character concepts, generics and gender. *Analytic Philosophy 56*(2), 111–141.

Leslie, S.-J. (2015b). The original sin of cognition: fear, prejudice and generalization. *The Journal of Philosophy*, forthcoming.

Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge, MA: The MIT Press.

Reuter, K. and G. Del Pinal (2016). Gender, social roles, and the normative dimension of conceptual representations. manuscript.

Rosch, E. (1999). Principles of categorization. In E. Margolis and S. Laurence (Eds.), *Concepts: Core Readings*, Chapter 8, pp. 189–206. Cambridge, MA: The MIT Press.

Rosch, E. (2011). "Slow lettuce": categories, concepts, fuzzy sets, and logical deduction. In R. Belohlavek and G. J. Klir (Eds.), *Concepts and Fuzzy Logic*, Chapter 4, pp. 89–120. Cambridge, MA: The MIT Press.

Roth, E. M. and E. J. Shoben (1983). The effect of context on the structure of categories. *Cognitive Psychology 15*(3), 346–378.

Sloman, S. A., B. C. Love, and W.-K. Ahn (1998). Feature centrality and conceptual coherence. *Cognitive Science 22*(2), 189–228.

Smith, E. E., D. L. Medin, L. J. Rips, and M. Keane (1988). Combining prototypes: a selective modification model. *Cognitive Science 12*, 485–527.

Strevens, M. (2000). The essentialist aspect of naive theories. *Cognition 74*, 149–175.